

Central Limit Theorem Analogues for Multicolour Urn Models

Noela S. Müller
Institute for Mathematics
J.W. Goethe University
60054 Frankfurt am Main
Germany

Email: nmueller@math.uni-frankfurt.de

April 4, 2017

Abstract

We prove a central limit theorem for the fine fluctuations of urn models under general assumptions on their generating matrices and initial configurations. The proof uses martingale techniques, extending a result of R.T. Smythe from 1995. This covers old results and gives some new results as well, including the m -ary search tree and B-urns.

MSC2010: 60F05, 60F15, 60C05, 60J10.

Keywords: Generalised Pólya urn models, Cyclic urn, m -ary search tree, Weak convergence, CLT analogue, Martingale central limit theorems.

1 Introduction and main result

1.1 Introduction

Consider a (generalised) Pólya urn scheme with $q \in \mathbb{N}$, $q \geq 2$, different colours that evolves in discrete time. At time zero, the beginning of the evolution, there is at least one ball in the urn. At each integer time, a ball is drawn from the urn uniformly at random and independently of the previous draws. Depending on its colour, one of q possible deterministic changes in the urn's ball configuration is realised. These changes are such that some balls are added to the urn and others are taken from the urn, with the restriction that only balls of the drawn colour can be removed from the urn at each step. The last condition, which might appear asymmetric at a first glance, is imposed so that one does not have to know the whole evolution of the urn process in order to know if, at a given instance, there actually are balls of the colour we wish to remove from the urn. We also assume that in total, we add the same (deterministic and positive) number of balls at each discrete time step such that the number of balls in the urn steadily grows to infinity as time increases. Under these premises, after a large number of draws, the relative frequencies of the numbers of balls of each colour will almost surely stabilise at some random or deterministic number. Our assumptions on the ball addition rules will be made precise below.

The aim of this paper is to derive a central limit theorem for the urn composition vector, i.e. for the number of balls of each colour. As we will see in Section 2, this vector can be interpreted as a sum over components of different sizes, roughly classified as *big* and *small* components. In order to treat these components on a common scale \sqrt{n} , the composition vector has to be centered in an appropriate way. This centering involves the deterministic expectation as well as random terms, if there are big components. See Theorem 1.3 below. Although this result also applies to urns

without big components, our main motivation stems from the study of so-called big urns, as there are plenty of results on the small case. For the special case of cyclic urns (see Example 2.1), a central limit theorem is derived in [16] by means of the contraction method. Results of this kind have also been proved in [12].

The proof of Theorem 1.3 benefits from a martingale point of view as taken in [18] and extends the result in [20] from 1995, using ideas and techniques from both papers. Urn models have been widely studied in [8], [10], [14] and [18], for example. Results closely related to the one given in this text can be found in [1], [2] and [12].

Notation. For a complex number $z \in \mathbb{C}$, we denote by $\Re(z)$, $\Im(z)$ and $|z|$ its real part, imaginary part and complex modulus, respectively. For a complex vector $v \in \mathbb{C}^q$ and $i \in \{1, \dots, q\}$, we denote by $v^{(i)}$ its i -th component and by v^t and v^* its transpose and conjugate transpose, respectively. Further let $|v|$ denote its L^1 -norm. We equip \mathbb{C}^q with the standard inner product $\langle \cdot, \cdot \rangle$, where $\langle u, v \rangle := u^* v$. If A is a subset of $\{1, \dots, q\}$, let δ_A denote the Dirac measure in A on $\{1, \dots, q\}$. Let $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ and $\mathbb{Z}_- := \{0, -1, -2, \dots\}$ denote the set of non-positive integers. We use Bachmann-Landau symbols in asymptotic statements. Finally, convergence in distribution is denoted as $\xrightarrow{\mathcal{L}}$.

1.2 Formalisation

We consider a generalised Pólya urn scheme, which evolves according to a discrete time Markov process, cf. [8] or [14]. At each time $n \geq 0$, there are balls of $q \in \mathbb{N}$ possible colours in the urn. By $X_n \in \mathbb{N}_0^q$, we denote the composition vector of the urn at time n , i.e.

$$X_n = \begin{pmatrix} X_n^{(1)} \\ X_n^{(2)} \\ \vdots \\ X_n^{(q-1)} \\ X_n^{(q)} \end{pmatrix},$$

where the component $X_n^{(j)}$ keeps track of the number of balls of colour j in the urn after n draws. The dynamics of the process are as follows:

At time 0, there are $X_0^{(j)}$ balls of colour j in the urn, where X_0 is a *deterministic* vector throughout the paper. Let the increments $\Delta_1, \dots, \Delta_q \in \mathbb{Z}^q$ be deterministic vectors and set $R := (\Delta_1, \dots, \Delta_q) \in \mathbb{Z}^{q \times q}$ to be the matrix with columns $\Delta_1, \dots, \Delta_q$. R will be called the generating matrix of the process. Note that R is the matrix transpose of the so-called replacement matrix. Immediately before each time $n \geq 1$, a ball is drawn uniformly at random and independently of all previous draws from the urn. If the ball drawn has colour i , then we add $\Delta_i^{(j)}$ balls of colour j for $j = 1, \dots, q$ to the urn. Thus, Δ_i describes the change in the urn composition if a ball of colour i is drawn.

The dynamics of the process are fully described by R and X_0 . Our basic assumptions are:

(A1) R is diagonalisable over \mathbb{C} .

(A2) R has constant column sum r .

(A3) $R_{i,j} \geq 0$ for $i \neq j$ and if $R_{i,i} < 0$, then $|R_{i,i}|$ divides $X_0^{(i)}$ and $R_{i,j}$ for all $1 \leq j \leq q$.

These assumptions are satisfied in most applications. (A2) guarantees a steady linear growth, while (A3) assures that the process does not get stuck by asking for an impossible removal of balls. Matrices with non-negative off-diagonal entries are called Metzler-Leontief matrices.

For more structural overview, we write $i \rightarrow j$ and say that colour i leads to colour j if, starting

with one ball of colour i , we have $\mathbb{P}(X_n^{(j)} > 0) > 0$ for some $n \in \mathbb{N}_0$. Equivalently, $(R^n)_{j,i} > 0$. We say that i and j communicate and write $i \leftrightarrow j$ if $i \rightarrow j$ and $j \rightarrow i$. The equivalence relation \leftrightarrow partitions the set $\{1, \dots, q\}$ of colours into equivalence classes $\mathcal{C}_1, \dots, \mathcal{C}_d$. If $d = 1$, the process is called *irreducible*. We will use the following spectral properties of irreducible Metzler-Leontief matrices, taken from [19] and presented here as in [6]:

Theorem 1.1. *Let $B = (B_{i,j})$ be an irreducible Metzler-Leontief matrix. Then, there exists a dominant eigenvalue τ of B such that*

- (i) τ is real, has multiplicity 1 and the associated left and right eigenvectors are positive;
- (ii) $\tau > \Re(\lambda)$ where $\lambda \neq \tau$ is any eigenvalue of B ;
- (iii) $\min_i \sum_j B_{i,j} \leq \tau \leq \max_i \sum_j B_{i,j}$;
- (iv) If there exists a non-negative vector x and a real number ρ such that $Bx \leq \rho x$, then $\rho \geq \tau$; $\rho = \tau$ if and only if $Bx = \rho x$;
- (v) $\sum_j B_{i,j} = 1$ for all i implies $\tau = 1$; and
- (vi) $\sum_j B_{i,j} \leq 1$ for all i with at least one strict inequality implies $\tau < 1$.

We write $\mathcal{C}_i \rightarrow \mathcal{C}_j$ and say that class \mathcal{C}_i leads to class \mathcal{C}_j if some (then all) colours in \mathcal{C}_i lead to some (then all) colours in \mathcal{C}_j . We call a class *dominant* if it does not lead to any other class except possibly itself. The classes $\mathcal{C}_1, \dots, \mathcal{C}_d$ decompose into three different types. Class \mathcal{C}_i is of type 1 if it is dominant and there is no $j \neq i$ with $\mathcal{C}_j \rightarrow \mathcal{C}_i$. Any dominant class which is not of type 1 is of type 2. All non dominant classes are of type 3. Subsequently, we may assume that the classes are ordered as follows: with $a, b, c \geq 0$ and $a + b + c = d$, classes $\mathcal{C}_1, \dots, \mathcal{C}_a$ are of type 1, classes $\mathcal{C}_{a+1}, \dots, \mathcal{C}_{a+c}$ are of type 2, and the remaining classes $\mathcal{C}_{a+c+1}, \dots, \mathcal{C}_d$ are of the third type. $\mathcal{C}_{a+c+1}, \dots, \mathcal{C}_d$ are ordered such that $\mathcal{C}_i \rightarrow \mathcal{C}_j$ implies $i \leq j$. Note that $a + c \geq 1$. The matrix R has the following lower triangular block structure:

$$R = \begin{pmatrix} T_{1,1} & & & & & \\ & \ddots & & & & \\ & & T_{a,a} & & & \\ & & & P_{1,1} & & \\ & & & * & \ddots & \\ & & & * & * & P_{b,b} \\ & & & * & * & * & Q_{1,1} \\ & & & * & * & * & & \ddots \\ & & & * & * & * & & & Q_{c,c} \end{pmatrix},$$

All blocks in R are irreducible Metzler-Leontief matrices. Blocks $T_{1,1}, \dots, T_{a,a}$ correspond to type 1 classes $\mathcal{C}_1, \dots, \mathcal{C}_a$, blocks $Q_{1,1}, \dots, Q_{c,c}$ to type 2 classes $\mathcal{C}_{a+1}, \dots, \mathcal{C}_{a+c}$ and blocks $P_{1,1}, \dots, P_{b,b}$ to type 3 classes $\mathcal{C}_{a+c+1}, \dots, \mathcal{C}_d$. The middle part corresponding to type 3 classes is a lower triangular block matrix in which beneath each of the blocks $P_{1,1}, \dots, P_{b,b}$, there is at least one positive entry. Similarly, to the left of each block $Q_{1,1}, \dots, Q_{c,c}$, there is at least one positive entry.

As the columns of $T_{1,1}, \dots, T_{a,a}$, $Q_{1,1}, \dots, Q_{c,c}$ sum to r , and the columns of $P_{1,1}, \dots, P_{b,b}$ sum to less than r , we can order the q eigenvalues of R by $r = \lambda_1 = \dots = \lambda_{a+c} > \Re(\lambda_{a+c+1}) \geq \dots \geq \Re(\lambda_q)$. Non dominant eigenvalues with equal real part are ordered by decreasing size of imaginary parts. If eigenvalue λ has multiplicity $m > 1$, λ is repeated m times in this list.

To avoid reduction to smaller urns, we will also assume

(A4) For each matrix $T_{1,1}, \dots, T_{a,a}$, $Q_{1,1}, \dots, Q_{c,c}$, no two columns are identical.

(A5) The initial composition of the urn is such that for all colours j , there exists $n \in \mathbb{N}_0$ with $\mathbb{P}\left(X_n^{(j)} > 0\right) > 0$.

In particular, we start with at least one ball from each class of type 1.

The spectrum of R . As R is diagonalisable, so are all blocks on its diagonal. We say that an eigenvalue λ_k belongs to class \mathcal{C}_m if it is an eigenvalue of the restriction of R to \mathcal{C}_m . Eigenvectors of blocks extend to eigenvectors of R in the following way:

1. If λ is an eigenvalue with multiplicity m of $T_{i,i}$ for some $1 \leq i \leq a$, then there exist m corresponding left (and right) eigenvectors which are zero on every colour outside $T_{i,i}$.
2. If λ is an eigenvalue with multiplicity m of $P_{i,i}$ for some $1 \leq i \leq b$, then there exist m corresponding left eigenvectors which are zero on colours in type 1 and type 2 classes and m right eigenvectors which are zero on type 1 classes.
3. Similarly, if λ is an eigenvalue with multiplicity m of $Q_{i,i}$ for some $1 \leq i \leq c$, then there exist m corresponding left eigenvectors which are zero on all colours in type 1 blocks, in type 2 blocks $Q_{j,j}$ for $j \in \{1, \dots, c\} \setminus \{i\}$ and in type 3 blocks $P_{j,j}$ that do not lead to $Q_{i,i}$. There exist m corresponding right eigenvectors that are zero on every colour outside $Q_{i,i}$.

We choose dual bases $\{u_1, \dots, u_q\}$ and $\{v_1, \dots, v_q\}$ of left and right (column) eigenvectors of R , respectively, with the above and the following additional properties.

Assume that λ_k is a multiple eigenvalue of submatrix $T_{i,i}$ (or $Q_{i,i}$) with $\Im(\lambda_k) \geq 0$. Let V be the random vector defined in Theorem 1.2 and the paragraph below. We then choose corresponding left eigenvectors u_x, u_y for λ_k such that for all $x \neq y$,

$$\begin{aligned} \langle \Re(u_x), \Re(u_y) \rangle_V &:= \sum_{m=1}^q V^{(m)} \Re(u_x^{(m)}) \Re(u_y^{(m)}) = 0, \\ \langle \Im(u_x), \Im(u_y) \rangle_V &:= \sum_{m=1}^q V^{(m)} \Im(u_x^{(m)}) \Im(u_y^{(m)}) = 0, \\ \langle \Re(u_x), \Im(u_y) \rangle_V &:= \sum_{m=1}^q V^{(m)} \Re(u_x^{(m)}) \Im(u_y^{(m)}) = 0. \end{aligned}$$

Note that this is possible as the entries of V are strictly positive on $T_{i,i}$ (or $Q_{i,i}$) with probability one. We do not put restrictions on basis eigenvectors of multiple eigenvalues in non-dominant classes. Furthermore, because u_x has non-zero components on exactly one dominant class (and maybe on some non-dominant components), multiple eigenvalues from different dominant classes are also “orthogonal” with respect to $\langle \cdot, \cdot \rangle_V$.

We assume that if λ_k is a complex eigenvalue with corresponding eigenvectors u_k, v_k , then the eigenvectors corresponding to $\bar{\lambda}_k$ are \bar{u}_k, \bar{v}_k . If λ_k is a real eigenvalue, then both u_k, v_k are chosen real. For $A \subseteq \{1, \dots, q\}$ and $v \in \mathbb{C}$, let v_A be the q dimensional vector defined by $v_A^{(i)} = v^{(i)} \cdot \delta_A(i)$. Let $\mathbf{1}$ denote the q dimensional all ones vector. We further assume that both left and right eigenvectors to r are real and are of the following form: $u_i := \mathbf{1}_{\mathcal{C}_i}$ for $i = 1, \dots, a$. Further, we can choose the remaining eigenvectors u_{a+1}, \dots, u_{a+c} orthogonal in such a way that $u_{a+s} = \mathbf{1}_{\mathcal{C}_{a+b+s}} + v_s$, where v_s is only nonzero on colour classes of type 3 leading to \mathcal{C}_{a+b+s} .

If R is irreducible, then $u_1 = (1, \dots, 1)^t$ is the only eigenvalue corresponding to r .

Having fixed the particular choice of eigenvectors, we turn to the spectral decomposition of \mathbb{C}^q relative to R . Let $\pi_k : \mathbb{C}^q \rightarrow \mathbb{C}$ be the linear map defined by

$$\pi_k(v) := u_k^* v.$$

Then $\pi_k(v)$ is the coefficient of the vector v_k in the representation of v with respect to the eigenvector basis $\{v_1, \dots, v_q\}$, i.e.

$$v = \sum_{k=1}^q \pi_k(v) v_k.$$

1.3 Convergence of proportions and main result

Remember that there are $|X_0|$ balls in the urn at time 0. As the urns we consider are balanced, there are $rn + |X_0|$ balls in the urn at time n .

For $n \geq 0$, let

$$Y_n := X_n - \mathbb{E}[X_n], \quad \mathcal{F}_n := \sigma(X_0, \dots, X_n).$$

We denote by $\text{Id}_{\mathbb{C}^q}$ the $q \times q$ identity matrix.

A first step towards a central limit theorem for urn models can be made by looking at the asymptotics of the proportions $X_n/(rn + |X_0|)$ of balls of different colours. In doing so, we first consider the case where $c = 1$, which is treated in Theorem 3.1 of [6]:

Theorem 1.2. *In the setting introduced above, suppose that (A1) - (A5) hold and assume $c = 1$. Then, as $n \rightarrow \infty$,*

$$\frac{X_n}{rn + |X_0|} \rightarrow \sum_{i=1}^a D^{(i)} v_i + D^{(a+1)} v_{a+1} =: V$$

almost surely, where $(D^{(1)}, \dots, D^{(a)}, D^{(a+1)})^t$ is a random vector with Dirichlet density

$$\Gamma \left(\sum_{j=1}^{a+1} \theta_j \right) \prod_{j=1}^{a+1} \frac{y_j^{\theta_j-1}}{\Gamma(\theta_j)} \mathbf{1}_{\{0 < y_1, \dots, y_{a+1} < 1, y_1 + \dots + y_{a+1} = 1\}} ((y_1, \dots, y_{a+1})^t).$$

Here, $\theta_j = |(X_0)_{C_j}|/r$, $j = 1, \dots, a$, and $\theta_{a+1} = |X_0|/r - \sum_{j=1}^a \theta_j$.

For the same result without identification of the Dirichlet components, see Theorem 3.5 in [18].

In the more general case $c \in \mathbb{N}$, the above result makes it plausible that

$$\frac{X_n}{rn + |X_0|} \rightarrow \sum_{i=1}^a D^{(i)} v_i + D^{(a+1)} (\Gamma_{a+1} v_{a+1} + \dots + \Gamma_{a+c} v_{a+c}) =: V$$

almost surely, where $(D^{(1)}, \dots, D^{(a)}, D^{(a+1)})^t$ is Dirichlet distributed with parameters as in Theorem 1.2. $\Gamma_{a+1}, \dots, \Gamma_{a+c}$ are random variables that sum to 1 almost surely and are independent of the Dirichlet random vector. Their distribution, which is an interesting question in its own right, is not specified, as they arise as martingale limits. A proof of this result can easily be obtained along the lines of the proofs given in the next section, and we omit the details here. In the following, the random vector V is used to denote the almost sure limit of the proportions $\frac{X_n}{rn + |X_0|}$. Note that it is zero in all type 3 components.

A possible view on Theorem 1.2 and its extension is to interpret the dynamics of the urn as a superposition of the dynamics of a classical Pólya urn and several irreducible urns. On a global scale, the isolated components $\mathcal{C}_1, \dots, \mathcal{C}_a, \mathcal{C}_{a+1} \cup \dots \cup \mathcal{C}_d$ of the process can be seen as the balls of a Pólya urn. Consequently, the asymptotic proportions among these supercolours are Dirichlet distributed. On an intermediate scale, the random variables $\Gamma_{a+1}, \dots, \Gamma_{a+c}$ are the asymptotic proportions of the non-isolated dominant classes inside supercolour $\mathcal{C}_{a+1} \cup \dots \cup \mathcal{C}_d$. On a local scale, inside a particular dominant component, the asymptotic proportions of balls are deterministic and given by the components of the right eigenvector coresponding to the class.

To formulate the main result of the paper, one further definition is needed. Set

$$\mathbf{M} := (\Re(\mathbf{v}_1), -\Im(\mathbf{v}_1), \Re(\mathbf{v}_2), -\Im(\mathbf{v}_2), \dots, \Re(\mathbf{v}_q), -\Im(\mathbf{v}_q)) \in \mathbb{R}^{q \times 2q}.$$

With this definition, the purpose of the paper is to prove and illustrate the following result.

Theorem 1.3. *In the setting above, suppose that (A1) - (A5) hold. Let $\mathfrak{p} := \max\{k : \Re(\lambda_k)/r > 1/2\}$.*

1. *Suppose that for all $k \in \{1, \dots, q\}$, $\Re(\lambda_k) \neq r/2$ for all λ_k that belong to a dominant class and that there are at least two colours in the dominant classes. Then there exist complex valued mean zero random variables $\Xi_1, \dots, \Xi_{\mathfrak{p}}$ such that*

$$\frac{1}{\sqrt{n}} \left(Y_n - \sum_{k=1}^{\mathfrak{p}} n^{\frac{\lambda_k}{r}} \Xi_k \mathbf{v}_k \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A_V) \quad (1)$$

as $n \rightarrow \infty$, where \mathcal{N} has a non degenerated, centered multivariate Gaussian mixture distribution with mixture components $\mathbf{V}^{(1)}, \dots, \mathbf{V}^{(q)}$ and covariance matrix

$$A_V := \mathbf{M} \Sigma_V \mathbf{M}^t,$$

where Σ_V is defined in (8) to (11) below Theorem 2.1. Furthermore, $(A_V)_{i,i} > 0$ almost surely for dominant colours i , whereas $(A_V)_{i,i} = 0$ almost surely for non dominant colours i .

2. *Suppose that there is some $k \in \{1, \dots, q\}$ such that $\Re(\lambda_k) = r/2$ and that λ_k belongs to a dominant class. Then there exist complex valued mean zero random variables $\Xi_1, \dots, \Xi_{\mathfrak{p}}$ such that*

$$\frac{1}{\sqrt{n \log(n)}} \left(Y_n - \sum_{k=1}^{\mathfrak{p}} n^{\frac{\lambda_k}{r}} \Xi_k \mathbf{v}_k \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A_V) \quad (2)$$

as $n \rightarrow \infty$, where \mathcal{N} has a non degenerated, centered multivariate Gaussian mixture distribution with mixture components $\mathbf{V}^{(1)}, \dots, \mathbf{V}^{(q)}$ and covariance matrix

$$A_V := \mathbf{M} \Sigma_V \mathbf{M}^t,$$

where Σ_V is defined in (12) below Theorem 2.1. $(A_V)_{i,i} > 0$ almost surely for dominant colours i that belong to the irreducible classes of eigenvalues with real part $r/2$, whereas $(A_V)_{i,i} = 0$ almost surely for all other colours.

Remark 1. Note that Theorem 1.3 covers all three classically distinguished cases where the rescaled composition vector is asymptotically normally distributed (e.g., m -ary search tree for $m \leq 26$), where it converges almost surely to some random limit (e.g., Pólya Urn) and where it exhibits some almost sure oscillating behaviour (e.g., m -ary search tree for $m \geq 27$ or cyclic urn for $m \geq 7$ colours).

Remark 2. The case where R is irreducible (or r is simple) and any other eigenvalue λ_k satisfies $\Re(\lambda_k) \leq r/2$ is known and treated in [8] and [20].

Acknowledgements. First of all, I thank Henning Sulzbach for many discussions and comments that led to a considerable improvement of the text. I am also indebted to Ralph Neininger, Kevin Leckey and Andrea Kuntschik for valuable remarks.

1.4 Projections and martingales

The key to the proof of Theorem 1.3 is an understanding of the magnitude and the asymptotic behaviour of certain components of X_n , namely the projection coefficients $\pi_k(Y_n)$. Most of the following facts are variants of known results and listed here to keep the text as self-contained as possible.

We first note that

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \left(\text{Id}_{\mathbb{C}^q} + \frac{R}{rn + |X_0|} \right) X_n, \quad (3)$$

yielding a vector-valued martingale

$$\left(\prod_{j=N}^{n-1} \left(\text{Id}_{\mathbb{C}^q} + \frac{R}{rj + |X_0|} \right)^{-1} X_n \right)_{n \geq N}$$

for some $N \in \mathbb{N}$ sufficiently large such that the occurring matrix inverses exist. This observation can be found below Definition 2.1 in [18].

The particular form of $\mathbb{E}[X_{n+1}|\mathcal{F}_n]$ leads to complex-valued martingales via projections on the eigenspaces of R . This idea is implicit in the work of Smythe [20] and more explicit in the proof of Theorem 3.5 in [18] for certain projections. We adopt it for all eigenspace projections.

Lemma 1.1 (Projection martingales). *(i) Let $k \in \{1, \dots, q\}$ be such that λ_k satisfies $\lambda_k + |X_0| \notin r\mathbb{Z}_-$. Define*

$$\gamma_n^{(k)} := \prod_{j=0}^{n-1} \left(1 + \frac{\bar{\lambda}_k}{rj + |X_0|} \right),$$

which is different from zero for all $n \geq 0$. Then

$$\mathbb{E}[\pi_k(X_n)] = \gamma_n^{(k)} \pi_k(X_0) = \frac{\Gamma\left(\frac{|X_0|}{r}\right) \pi_k(X_0)}{\Gamma\left(\frac{|X_0| + \bar{\lambda}_k}{r}\right)} \cdot n^{\frac{\bar{\lambda}_k}{r}} + O\left(n^{\Re\left(\frac{\bar{\lambda}_k}{r}\right) - 1}\right),$$

as $n \rightarrow \infty$. $(M_n^{(k)})_{n \geq 0}$, defined by

$$M_n^{(k)} := \left(\gamma_n^{(k)} \right)^{-1} \cdot \pi_k(Y_n),$$

is a complex-valued martingale with mean zero.

(ii) Let $k \in \{1, \dots, q\}$ be such that λ_k satisfies $\lambda_k + |X_0| \in r\mathbb{Z}_-$. Define

$$\gamma_n^{(k)} := \prod_{j=-\frac{\lambda_k + |X_0|}{r} + 1}^{n-1} \left(1 + \frac{\lambda_k}{rj + |X_0|} \right),$$

which is different from zero for all $n \geq -\frac{\lambda_k + |X_0|}{r} + 1$. Then

$$\mathbb{E}[\pi_k(X_n)] = 0,$$

for all $n \geq -\frac{\lambda_k + |X_0|}{r} + 1$. $(M_n^{(k)})_{n \geq -\frac{\lambda_k + |X_0|}{r} + 1}$, defined by

$$M_n^{(k)} := \left(\gamma_n^{(k)} \right)^{-1} \cdot \pi_k(Y_n),$$

is a complex-valued martingale with mean zero.

Proof. Let $k \in \{1, \dots, q\}$ and $n \geq 0$. As a direct consequence of (3) for all $n \geq 0$,

$$\mathbb{E}[\pi_k(X_{n+1})|\mathcal{F}_n] = \left(1 + \frac{\bar{\lambda}_k}{rn + |X_0|}\right) \pi_k(X_n)$$

almost surely and $\left(M_n^{(k)}\right)_n$ is a martingale in each of the two cases. In particular,

$$\mathbb{E}[\pi_k(X_n)] = \prod_{j=0}^{n-1} \left(1 + \frac{\bar{\lambda}_k}{rj + |X_0|}\right) \pi_k(X_0),$$

which is zero in the second case. In the first case, by Stirling's formula,

$$\gamma_n^{(k)} = \frac{\Gamma\left(\frac{|X_0|}{r}\right)}{\Gamma\left(\frac{|X_0| + \bar{\lambda}_k}{r}\right)} \cdot \frac{\Gamma\left(n + \frac{|X_0|}{r} + \frac{\bar{\lambda}_k}{r}\right)}{\Gamma\left(n + \frac{|X_0|}{r}\right)} = \frac{\Gamma\left(\frac{|X_0|}{r}\right)}{\Gamma\left(\frac{|X_0| + \bar{\lambda}_k}{r}\right)} \cdot n^{\frac{\bar{\lambda}_k}{r}} + O\left(n^{\Re\left(\frac{\bar{\lambda}_k}{r}\right) - 1}\right)$$

as $n \rightarrow \infty$. This implies the claim. \square

The martingales of the preceding proposition can be divided into two classes: convergent and non convergent. The corresponding eigenvalues are sometimes referred to as “big” and “small”, respectively. The remainder of this section will be devoted to properties of the convergent martingales and their limits.

Lemma 1.2 (Martingale limits). *For each $k \in \{1, \dots, q\}$ such that $\Re(\lambda_k) > r/2$, there exists a complex valued mean zero random variable Ξ_k such that*

$$M_n^{(k)} \rightarrow \frac{\Gamma(|X_0|/r + \bar{\lambda}_k/r)}{\Gamma(|X_0|/r)} \Xi_k$$

almost surely and in L^2 as $n \rightarrow \infty$.

Remark 3. The random variables Ξ_k in Theorem 1.3 and Lemma 1.2 are identical.

Proof. We apply the L^2 -martingale convergence theorem and show boundedness of second moments.

$$\mathbb{E}[|\pi_k(X_{n+1})|^2|\mathcal{F}_n] = \left(1 + \frac{2\Re(\lambda_k)}{rn + |X_0|}\right) |\pi_k(X_n)|^2 + \sum_{j=1}^q \frac{X_n^{(j)}}{rn + |X_0|} |\pi_k(\Delta_j)|^2.$$

Set $C_k := \sum_{j=1}^q |\pi_k(\Delta_j)|^2$. With this,

$$\mathbb{E}[|\pi_k(X_{n+1})|^2|\mathcal{F}_n] \leq \left(1 + \frac{2\Re(\lambda_k)}{rn + |X_0|}\right) |\pi_k(X_n)|^2 + C_k$$

and thus

$$\begin{aligned} \mathbb{E}[|\pi_k(X_n)|^2] &\leq \prod_{j=0}^{n-1} \left(1 + \frac{2\Re(\lambda_k)}{rj + |X_0|}\right) \mathbb{E}[|\pi_k(X_0)|^2] + C_k \prod_{j=1}^{n-1} \left(1 + \frac{2\Re(\lambda_k)}{rj + |X_0|}\right) \sum_{m=0}^{n-1} \prod_{j=1}^m \left(1 + \frac{2\Re(\lambda_k)}{rj + |X_0|}\right)^{-1} \\ &= \prod_{j=0}^{n-1} \left(1 + \frac{2\Re(\lambda_k)}{rj + |X_0|}\right) \left(\mathbb{E}[|\pi_k(X_0)|^2] + C_k \left(1 + \frac{2\Re(\lambda_k)}{|X_0|}\right)^{-1} \sum_{m=0}^{n-1} \prod_{j=1}^m \left(1 + \frac{2\Re(\lambda_k)}{rj + |X_0|}\right)^{-1} \right) \\ &= O\left(n^{2\Re(\lambda_k)/r}\right) \end{aligned}$$

as $n \rightarrow \infty$, because $\Re(\lambda_k) > r/2$. By Jensen's inequality, $|\mathbb{E}[\pi_k(X_n)]|^2 = O(n^{2\Re(\lambda_k)/r})$ as $n \rightarrow \infty$. Thus

$$\mathbb{E} \left[\left| M_n^{(k)} \right|^2 \right] = \left| \gamma_n^{(k)} \right|^{-2} \left(\mathbb{E} [|\pi_k(X_n)|^2] - |\mathbb{E} [\pi_k(X_n)]|^2 \right) = O(1)$$

as $n \rightarrow \infty$. By the L^2 -martingale convergence theorem, $M_n^{(k)}$ converges almost surely and in L^2 to some complex random variable which we normalise such that $\pi_k(Y_n) - n^{\bar{\lambda}_k/r} \Xi_k \rightarrow 0$ almost surely. \square

Remark 4. Recall the definition of the random proportions $D^{(1)}, \dots, D^{(a+1)}\Gamma_{a+1}, \dots, D^{(a+1)}\Gamma_{a+c}$ in Theorem 1.2 and below. For $k \in \{1, \dots, a\}$, we have the relation

$$\frac{\Xi_k}{r} = D^{(k)} - \frac{\pi_k(X_0)}{|X_0|}. \quad (4)$$

For $k \in \{a+1, \dots, a+c\}$,

$$\frac{\Xi_k}{r} = D^{(a+1)}\Gamma_k - \frac{\pi_k(X_0)}{|X_0|}.$$

This yields the representation

$$V = \sum_{k=1}^{a+c} \left(\frac{\Xi_k}{r} + \frac{\pi_k(X_0)}{|X_0|} \right) \cdot v_k \quad (5)$$

for the proportion vector V via the martingale limits of Lemma 1.2.

Also note that all random variables $D^{(1)}, \dots, D^{(a+1)}\Gamma_{a+1}, \dots, D^{(a+1)}\Gamma_{a+c}$ are strictly positive almost surely: This is clear for $D^{(1)}, \dots, D^{(a+1)}$. By condensing all type 2 colours to a single colour, Theorem 1.2 applies and it follows that the proportion of type 3 colours tends to zero almost surely (as was already stated in the first section). For $k \in \{a+1, \dots, a+c\}$, we see that

$$D^{(a+1)}\Gamma_k = \lim_{n \rightarrow \infty} \frac{\pi_k(X_n)}{rn + |X_0|} = \lim_{n \rightarrow \infty} \frac{|(X_n)_{C_k}|}{rn + |X_0|}$$

is the almost sure limit of the proportion of balls in class C_k . It can now easily be seen by a coupling argument with a classical Pólya urn, that these proportions are positive almost surely: Given an arbitrary initial configuration, let the urn evolve until there are balls of all dominant classes in the urn. Regard the ball composition at this time as the new initial configuration. From this time on, each time a ball of colour i in a type 3 class is drawn, instead of following the original rules, give all its children outside the class of i colour i , too. In the modified urn, type 3 classes become dominant, and almost surely, the proportions of colours in type 2 classes tend to a strictly positive limit. As there are at least as many type 2 balls in this urn as in the urn with unchanged colours, the claim follows.

Corollary 1.4 (Random limits). *Under conditions (A1) to (A5), Ξ_1, \dots, Ξ_{a+c} are almost surely non degenerated unless r is simple. In this case, $\Xi_1 = 0$. $\Xi_{a+c+1}, \dots, \Xi_p$ are almost surely non degenerated.*

Proof. Note that for the isolated type 1 limits Ξ_1, \dots, Ξ_a , the claim immediately follows from Theorem 1.2 and identity (4): For $k \in \{1, \dots, a\}$,

$$\text{Var}(\Xi_k) = r^2 \text{Var} \left(D^{(k)} \right),$$

where Dirichlet vector D has parameters $\frac{|(X_0)_{C_1}|}{r}, \dots, \frac{|(X_0)_{C_a}|}{r}, \frac{|X_0|}{r} - \sum_{j=1}^a \frac{|(X_0)_{C_j}|}{r}$.

More generally and without reference to Theorem 1.2, we can use orthogonality of martingale increments to see that for $k \in \{1, \dots, p\}$,

$$\begin{aligned}\mathbb{E} [|\Xi_k|^2] &= \mathbb{E} \left[\left| \Xi_k - M_0^{(k)} \right|^2 \right] = \sum_{j=0}^{\infty} \mathbb{E} \left[\left| M_{j+1}^{(k)} - M_j^{(k)} \right|^2 \right] \\ &= \sum_{j=0}^{\infty} \left| \gamma_{j+1}^{(k)} \right|^{-2} \mathbb{E} \left[\left| \pi_k(X_{j+1} - X_j) - \frac{\bar{\lambda}_k}{rj + |X_0|} \pi_k(X_j) \right|^2 \right].\end{aligned}$$

The requirement of zero variance completely determines the evolution of projection k in each draw: The expression on the right hand side is only equal to zero if for all $j \geq 0$,

$$\pi_k(X_{j+1} - X_j) = \frac{\bar{\lambda}_k}{rj + |X_0|} \pi_k(X_j) \quad (6)$$

almost surely (note that $\lambda_k \neq 0$). This in particular means that the value of $\pi_k(X_{j+1} - X_j)$ is independent of the colour of the $(j+1)$ -th ball drawn from the urn. We will see that this is not possible under our assumptions.

First assume that there is an initial configuration X_0 that is compatible with (A1) to (A5) and has $\pi_k(X_0) = 0$. Under this initial configuration, $\pi_k(X_j) = 0$ for all $j \geq 0$ almost surely because of (6), and with probability one,

$$0 = \pi_k(X_{j+1} - X_j) = \bar{\lambda}_k \bar{u}_k^{(N_{j+1})}$$

for all $j \geq 0$. Here, N_{j+1} denotes the colour seen in the $(j+1)$ st draw. For each colour $f \in \{1, \dots, q\}$, there is $n \in \mathbb{N}_0$ with $\mathbb{P} \left(X_n^{(f)} > 0 \right) > 0$, due to assumption (A5). The last equation yields that $u_k^{(f)} = 0$. So $u_k = 0$, which is a contradiction.

The last paragraph showed that there is no admissible choice of initial configuration such that $\pi_k(X_0) = 0$. For $\pi_k(X_0) \neq 0$, $\pi_k(X_{j+1} - X_j) \neq 0$ for all j almost surely because of (6). If there is more than one dominant colour or if $\bar{\lambda}_k$ belongs to a type 3 class, it follows immediately from our choice of left eigenvectors that $\pi_k(X_{j+1} - X_j) \neq 0$ for all j almost surely is not possible: According to the mechanism of the urn, once there is a ball of a dominant class in the urn, there are balls of its class in the urn at all future times. So almost surely, there is a time N at which there are balls of all dominant classes in the urn. From this point on, there is a positive probability of drawing balls that lead to no change in the class under consideration. Now assume that there is only one dominant class and that λ_k belongs to this class. As the proportions of balls in this class converge to a positive limit with probability one, there is a time N from which on there are balls of each colour of the dominant class in the urn. This implies that $u_k^{(i)} = u_k^{(j)}$ for all colours i, j in this class. So $\lambda_k = r$, and $u_k = (1, \dots, 1)^t$ is the only projection that induces a deterministic limit. \square

Lemma 1.3 (Speed of convergence). *Let $k \in \{1, \dots, q\}$ be such that $\Re(\lambda_k) > r/2$. Then*

$$\left\| \frac{\Gamma(|X_0|/r + \bar{\lambda}_k/r)}{\Gamma(|X_0|/r)} \Xi_k - M_n^{(k)} \right\|_{L^2} = O \left(n^{1/2 - \Re(\lambda_k)/r} \right) \quad (7)$$

as $n \rightarrow \infty$.

Proof. We use the decomposition

$$\begin{aligned}
\left\| \frac{\Gamma(|X_0|/r + \bar{\lambda}_k/r)}{\Gamma(|X_0|/r)} \Xi_k - M_n^{(k)} \right\|_{L^2}^2 &= \sum_{j=n}^{\infty} \mathbb{E} \left[\left| M_{j+1}^{(k)} - M_j^{(k)} \right|^2 \right] \\
&= \sum_{j=n}^{\infty} \left| \gamma_{j+1}^{(k)} \right|^{-2} \mathbb{E} \left[\left| \pi_k(X_{j+1} - X_j) - \frac{\bar{\lambda}_k}{rj + |X_0|} \pi_k(X_j) \right|^2 \right] \\
&= \sum_{j=n}^{\infty} \left| \gamma_{j+1}^{(k)} \right|^{-2} \left(\mathbb{E} [|\pi_k(X_{j+1} - X_j)|^2] - \left| \frac{\lambda_k}{rj + |X_0|} \right|^2 \mathbb{E} [|\pi_k(X_j)|^2] \right) \\
&\leq \sum_{j=n}^{\infty} \left| \gamma_{j+1}^{(k)} \right|^{-2} \mathbb{E} [|\pi_k(X_{j+1} - X_j)|^2] \\
&\leq Cn^{1-2\Re(\lambda_k)/r}
\end{aligned}$$

as $|\pi_k(X_{j+1} - X_j)|^2$ can only take q values, independently of j . \square

2 Proof of Theorem 1.3

Let $p := \max\{k : \Re(\lambda_k)/r > 1/2\}$ as in Theorem 1.3. We now study the joint fluctuations of all projections. There are two cases as in Theorem 1.3. If there is no k such that $\Re(\lambda_k)/r = 1/2$, let

$$Z_n := \frac{1}{\sqrt{n}} \begin{pmatrix} \Re(\pi_1(Y_n) - \gamma_n^{(1)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_1/r)}{\Gamma(|X_0|/r)} \Xi_1) \\ \Im(\pi_1(Y_n) - \gamma_n^{(1)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_1/r)}{\Gamma(|X_0|/r)} \Xi_1) \\ \Re(\pi_2(Y_n) - \gamma_n^{(2)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_2/r)}{\Gamma(|X_0|/r)} \Xi_2) \\ \Im(\pi_2(Y_n) - \gamma_n^{(2)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_2/r)}{\Gamma(|X_0|/r)} \Xi_2) \\ \vdots \\ \Re(\pi_p(Y_n) - \gamma_n^{(p)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_p/r)}{\Gamma(|X_0|/r)} \Xi_p) \\ \Im(\pi_p(Y_n) - \gamma_n^{(p)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_p/r)}{\Gamma(|X_0|/r)} \Xi_p) \\ \Re(\pi_{p+1}(Y_n)) \\ \Im(\pi_{p+1}(Y_n)) \\ \vdots \\ \Re(\pi_q(Y_n)) \\ \Im(\pi_q(Y_n)) \end{pmatrix}.$$

If there is such a k , let

$$Z_n := \frac{1}{\sqrt{n \log(n)}} \begin{pmatrix} \Re(\pi_1(Y_n) - \gamma_n^{(1)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_1/r)}{\Gamma(|X_0|/r)} \Xi_1) \\ \Im(\pi_1(Y_n) - \gamma_n^{(1)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_1/r)}{\Gamma(|X_0|/r)} \Xi_1) \\ \Re(\pi_2(Y_n) - \gamma_n^{(2)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_2/r)}{\Gamma(|X_0|/r)} \Xi_2) \\ \Im(\pi_2(Y_n) - \gamma_n^{(2)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_2/r)}{\Gamma(|X_0|/r)} \Xi_2) \\ \vdots \\ \Re(\pi_p(Y_n) - \gamma_n^{(p)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_p/r)}{\Gamma(|X_0|/r)} \Xi_p) \\ \Im(\pi_p(Y_n) - \gamma_n^{(p)} \frac{\Gamma(|X_0|/r + \bar{\lambda}_p/r)}{\Gamma(|X_0|/r)} \Xi_p) \\ \Re(\pi_{p+1}(Y_n)) \\ \Im(\pi_{p+1}(Y_n)) \\ \vdots \\ \Re(\pi_q(Y_n)) \\ \Im(\pi_q(Y_n)) \end{pmatrix}.$$

Note that some components of Z_n may be equal or 0. The aim of the next section is to show convergence in distribution of Z_n to a multivariate mixed Gaussian distribution.

Theorem 2.1. *As $n \rightarrow \infty$,*

$$Z_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_V),$$

where $\mathcal{N}(0, \Sigma_V)$ is a mixed Gaussian random vector with mixture components $V^{(1)}, \dots, V^{(q)}$ and covariance matrix Σ_V defined in (8) to (11) and (12).

Definition of Σ_V . There are two cases. First assume that for all eigenvalues of dominant classes, $\Re(\lambda_k) \neq r/2$. The possible *non-zero* entries of the $2q \times 2q$ matrix Σ_V are then given by

$$(\Sigma_V)_{2k-1, 2\ell-1} := \begin{cases} r^2 \left(\frac{\Xi_k}{r} + \frac{\pi_k(X_0)}{|X_0|} \right) \left(1 - \left(\frac{\Xi_k}{r} + \frac{\pi_k(X_0)}{|X_0|} \right) \right), & k = \ell, \lambda_k = r \\ -r^2 \left(\frac{\Xi_k}{r} + \frac{\pi_k(X_0)}{|X_0|} \right) \left(\frac{\Xi_\ell}{r} + \frac{\pi_\ell(X_0)}{|X_0|} \right), & k \neq \ell, \lambda_k = \lambda_\ell = r \\ \sum_{m=1}^q V^{(m)} \Re \left(\frac{\left(\frac{\lambda_k + \lambda_\ell}{r} - 1 \right) \bar{\lambda}_k \bar{\lambda}_\ell \bar{u}_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\lambda_k + \lambda_\ell}{r} \right|^2} + \frac{\left(\frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} - 1 \right) \lambda_k \bar{\lambda}_\ell u_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} \right|^2} \right), & r/2 < \Re(\lambda_k), \Re(\lambda_\ell) < r \\ \sum_{m=1}^q V^{(m)} \Re \left(\frac{\left(1 - \frac{\lambda_k + \lambda_\ell}{r} \right) \bar{\lambda}_k \bar{\lambda}_\ell \bar{u}_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\lambda_k + \lambda_\ell}{r} \right|^2} + \frac{\left(1 - \frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} \right) \lambda_k \bar{\lambda}_\ell u_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} \right|^2} \right), & \Re(\lambda_k), \Re(\lambda_\ell) < r/2 \end{cases} \quad (8)$$

and

$$(\Sigma_V)_{2k, 2\ell} := \begin{cases} \sum_{m=1}^q V^{(m)} \Re \left(-\frac{\left(\frac{\lambda_k + \lambda_\ell}{r} - 1 \right) \bar{\lambda}_k \bar{\lambda}_\ell \bar{u}_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\lambda_k + \lambda_\ell}{r} \right|^2} + \frac{\left(\frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} - 1 \right) \lambda_k \bar{\lambda}_\ell u_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} \right|^2} \right), & r/2 < \Re(\lambda_k), \Re(\lambda_\ell) < r \\ \sum_{m=1}^q V^{(m)} \Re \left(-\frac{\left(1 - \frac{\lambda_k + \lambda_\ell}{r} \right) \bar{\lambda}_k \bar{\lambda}_\ell \bar{u}_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\lambda_k + \lambda_\ell}{r} \right|^2} + \frac{\left(1 - \frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} \right) \lambda_k \bar{\lambda}_\ell u_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} \right|^2} \right), & \Re(\lambda_k), \Re(\lambda_\ell) < r/2 \end{cases} \quad (9)$$

and

$$(\Sigma_V)_{2k-1, 2\ell} := \begin{cases} \sum_{m=1}^q V^{(m)} \Im \left(\frac{\left(\frac{\lambda_k + \lambda_\ell}{r} - 1 \right) \bar{\lambda}_k \bar{\lambda}_\ell \bar{u}_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\lambda_k + \lambda_\ell}{r} \right|^2} + \frac{\left(\frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} - 1 \right) \lambda_k \bar{\lambda}_\ell u_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} \right|^2} \right), & r/2 < \Re(\lambda_k), \Re(\lambda_\ell) < r \\ \sum_{m=1}^q V^{(m)} \Im \left(\frac{\left(1 - \frac{\lambda_k + \lambda_\ell}{r} \right) \bar{\lambda}_k \bar{\lambda}_\ell \bar{u}_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\lambda_k + \lambda_\ell}{r} \right|^2} + \frac{\left(1 - \frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} \right) \lambda_k \bar{\lambda}_\ell u_k^{(m)} \bar{u}_\ell^{(m)}}{2 \left| 1 - \frac{\bar{\lambda}_k + \bar{\lambda}_\ell}{r} \right|^2} \right), & \Re(\lambda_k), \Re(\lambda_\ell) < r/2 \end{cases} \quad (10)$$

as well as

$$(\Sigma_V)_{2k, 2\ell-1} := (\Sigma_V)_{2\ell-1, 2k}. \quad (11)$$

In the case where there is some dominant class such that $\Re(\lambda_k)/r = 1/2$, the matrix Σ_V has a lot more zero entries due to the scaling. Its non zero entries are in places $(2k-1, 2k-1)$ and $(2k, 2k)$ for dominant eigenvalues with $\Re(\lambda_k)/r = 1/2$. For these k ,

$$(\Sigma_V)_{2k-1, 2k-1} = (\Sigma_V)_{2k, 2k} = \frac{|\lambda_k|^2}{2} \sum_{m=1}^q V^{(m)} |u_k^{(m)}|^2. \quad (12)$$

Comments on the covariance structure Σ_V . Let us first consider the case where there is no dominant eigenvalue with real part $r/2$. In this case:

1. In the limit, $Z_n^{(i)}$ and $Z_n^{(j)}$ are independent for $i \in \{1, \dots, 2p\}$ and $j \in \{2p+1, \dots, 2q\}$.
2. The “proportion” components $Z_n^{(1)}, \dots, Z_n^{(2(a+c))}$ are independent of all other components.
3. The fluctuations of projections for eigenvalues λ_k corresponding to type 3 classes vanish in the \sqrt{n} scaling. This might be due to the fact that there are too little draws from these classes compared to the other classes. So Theorem 2.1 says nothing about the fluctuations within these classes (or, at least, nothing particularly interesting), as the draws from the dominant colours dominate in the limit and there is too little fluctuation among the remaining colours.

If there are dominant eigenvalues with real part $r/2$:

1. The fluctuations of the other projections are still of order \sqrt{n} and they tend to zero in the $\sqrt{n \log(n)}$ scaling.
2. Real and imaginary parts of the $\Re(\lambda_k) = r/2$ -components are independent, and they are also independent of each other.

Example 2.1 (Cyclic urn). Consider the cyclic urn with $q \geq 2$ colours, whose generating matrix is given by

$$R := \begin{pmatrix} 0 & 0 & 0 & \cdot & \cdot & 0 & 1 \\ 1 & 0 & 0 & \cdot & \cdot & 0 & 0 \\ 0 & 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & 1 & 0 \end{pmatrix}. \quad (13)$$

R is irreducible and has simple eigenvalues

$$1, \exp\left(\frac{2\pi i}{q}\right), \exp\left(\frac{2\pi i(q-1)}{q}\right), \dots, \exp\left(\frac{2\pi i \lfloor \frac{q}{2} \rfloor}{q}\right), \exp\left(\frac{2\pi i \lceil \frac{q}{2} \rceil}{q}\right),$$

if 2 does not divide q . These are the q -th roots of unity. If 2 divides q , the last two eigenvalues in the above list are replaced by a single -1 .

For the cyclic urn model, the covariances of the projections are calculated explicitly in [16]. They take a particularly simple form as the corresponding left eigenvectors can be chosen as an orthogonal basis. In particular, all components of Z_n are asymptotically independent, except for components corresponding to complex-conjugated eigenvalues. As R is irreducible, the “proportion” projection is deterministic and

$$(\Sigma_V)_{1,1} = (\Sigma_V)_{2,2} = 0$$

for all values of q .

To lighten the notation a bit, let us assume that q is a multiple of two in this example. The case that there is no eigenvalue with real part $1/2$ corresponds to the case $6 \nmid q$. In this case, the diagonal entries of Σ_V are given by

$$(\Sigma_V)_{2k-1,2k-1} = (\Sigma_V)_{2k,2k} = \frac{1}{2|2\lambda_k - 1|}$$

for $k \in \{2, \dots, q-1\}$ and

$$(\Sigma_V)_{q-1,q-1} = \frac{1}{3}, \quad (\Sigma_V)_{q,q} = 0.$$

All off-diagonal entries are zero, except if $\lambda_\ell = \bar{\lambda}_k$. Then

$$(\Sigma_V)_{2k-1, 2\ell-1} = \frac{1}{2|2\lambda_k - 1|}, \quad (\Sigma_V)_{2k-1, 2\ell} = (\Sigma_V)_{2k, 2\ell-1} = 0, \quad (\Sigma_V)_{2k, 2\ell} = -\frac{1}{2|2\lambda_k - 1|}.$$

If $6 \mid q$, additionally all diagonal entries for $k \notin \{4 \cdot (\frac{q}{6} - 1) + 3, \dots, 4 \cdot \frac{q}{6} + 2\}$ are zero. For $k \in \{4 \cdot (\frac{q}{6} - 1) + 3, \dots, 4 \cdot \frac{q}{6} + 2\}$,

$$(\Sigma_V)_{k, k} = \frac{1}{2}.$$

Also, as complex conjugate projections have the same real part and imaginary part multiplied by -1 ,

$$(\Sigma_V)_{4 \cdot (\frac{q}{6} - 1) + 3, 4 \cdot \frac{q}{6} + 1} = \frac{1}{2}, \quad (\Sigma_V)_{4 \cdot \frac{q}{6}, 4 \cdot \frac{q}{6} + 2} = -\frac{1}{2}.$$

◇

In the proof of Theorem 2.1, we will employ Corollary 3.1 from the book [7] of P. Hall and C. Heyde:

Proposition 2.2. *Let $\{S_{n,j}, \mathcal{F}_{n,j}, 1 \leq j \leq k_n, n \geq 1\}$ be a zero-mean, square-integrable martingale array with increments $I_{n,j}$ and let η^2 be an a.s. finite random variable. Suppose that for all $\varepsilon > 0$,*

$$\sum_j \mathbb{E}[I_{n,j}^2 I(|I_{n,j}| > \varepsilon) | \mathcal{F}_{n,j-1}] \xrightarrow{\mathbb{P}} 0, \quad (14)$$

and

$$\sum_j \mathbb{E}[I_{n,j}^2 | \mathcal{F}_{n,j-1}] \xrightarrow{\mathbb{P}} \eta^2, \quad (15)$$

and $\mathcal{F}_{n,j} \subseteq \mathcal{F}_{n+1,j}$ for $1 \leq j \leq k_n$, $n \geq 1$. Then

$$S_{n, k_n} = \sum_j I_{n,j} \xrightarrow{d} Z,$$

where the random variable Z has characteristic function

$$\mathbb{E} \left[\exp \left(-\frac{1}{2} \eta^2 t^2 \right) \right].$$

Proof of Theorem 2.1. We will prove Theorem 2.1 via the Cramér-Wold device, thus proving weak convergence of all linear combinations of the components of Z_n instead. Let $\alpha_1, \dots, \alpha_{2q} \in \mathbb{R}$. Our aim is to apply Proposition 2.2 to an appropriate martingale array. To this end, in simple words, we decompose the sum $\alpha_1 Z_{n,1} + \dots + \alpha_{2q} Z_{n,q}$ into a sum of weighted martingale differences. We set $\mathcal{F}_{n,i} := \sigma(X_0, \dots, X_i)$. This definition satisfies the condition on the filtration from Proposition 2.2.

First consider the case where for all dominant eigenvalues, $\Re(\lambda_k) \neq r/2$.

We rewrite the given linear combination $\alpha_1 Z_n^{(1)} + \dots + \alpha_{2q} Z_n^{(2q)}$ as a sum of martingale differences. We simultaneously consider real and imaginary part of each eigenspace coefficient. For $1 \leq k \leq p$,

write

$$\begin{aligned}
& \alpha_{2k-1} Z_{n,2k-1} + \alpha_{2k} Z_{n,2k} \\
&= \frac{1}{\sqrt{n}} \sum_{j=n}^{\infty} \left(\alpha_{2k-1} \Re \left(\gamma_n^{(k)} \left(M_j^{(k)} - M_{j+1}^{(k)} \right) \right) + \alpha_{2k} \Im \left(\gamma_n^{(k)} \left(M_j^{(k)} - M_{j+1}^{(k)} \right) \right) \right) \\
&= \frac{1}{\sqrt{n}} (\alpha_{2k-1} \Re(\gamma_n^{(k)}) + \alpha_{2k} \Im(\gamma_n^{(k)})) \sum_{j=n}^{\infty} \Re \left(M_j^{(k)} - M_{j+1}^{(k)} \right) \\
&+ \frac{1}{\sqrt{n}} (\alpha_{2k} \Re(\gamma_n^{(k)}) - \alpha_{2k-1} \Im(\gamma_n^{(k)})) \sum_{j=n}^{\infty} \Im \left(M_j^{(k)} - M_{j+1}^{(k)} \right) \\
&=: \beta_{2k-1}(n) \sum_{j=n}^{\infty} \Re \left(M_j^{(k)} - M_{j+1}^{(k)} \right) + \beta_{2k}(n) \sum_{j=n}^{\infty} \Im \left(M_j^{(k)} - M_{j+1}^{(k)} \right).
\end{aligned}$$

Set $g := \max \left\{ -\frac{\lambda_k + |X_0|}{r} + 1 : \lambda_k + |X_0| \in r\mathbb{Z}_- \right\}$. Then for $p+1 \leq k \leq q$,

$$\begin{aligned}
& \alpha_{2k-1} Z_{n,2k-1} + \alpha_{2k} Z_{n,2k} \\
&= \frac{1}{\sqrt{n}} (\alpha_{2k-1} \Re(\gamma_n^{(k)}) + \alpha_{2k} \Im(\gamma_n^{(k)})) \sum_{j=g}^{n-1} \Re \left(M_{j+1}^{(k)} - M_j^{(k)} \right) \\
&+ \frac{1}{\sqrt{n}} (\alpha_{2k} \Re(\gamma_n^{(k)}) - \alpha_{2k-1} \Im(\gamma_n^{(k)})) \sum_{j=g}^{n-1} \Im \left(M_{j+1}^{(k)} - M_j^{(k)} \right) \\
&+ \frac{1}{\sqrt{n}} (\alpha_{2k-1} \Re(\gamma_n^{(k)}) + \alpha_{2k} \Im(\gamma_n^{(k)})) \Re \left(M_g^{(k)} \right) \\
&+ \frac{1}{\sqrt{n}} (\alpha_{2k} \Re(\gamma_n^{(k)}) - \alpha_{2k-1} \Im(\gamma_n^{(k)})) \Im \left(M_g^{(k)} \right) \\
&=: \beta_{2k-1}(n) \sum_{j=g}^{n-1} \Re \left(M_{j+1}^{(k)} - M_j^{(k)} \right) + \beta_{2k}(n) \sum_{j=g}^{n-1} \Im \left(M_{j+1}^{(k)} - M_j^{(k)} \right) + r_n(k).
\end{aligned}$$

So setting $r_n := \sum_{k=p+1}^q r_n(k)$,

$$\begin{aligned}
\sum_{k=1}^{2q} \alpha_k Z_{n,k} &= \sum_{k=1}^p \left(\beta_{2k-1}(n) \sum_{j=n}^{\infty} \Re \left(M_j^{(k)} - M_{j+1}^{(k)} \right) + \beta_{2k}(n) \sum_{j=n}^{\infty} \Im \left(M_j^{(k)} - M_{j+1}^{(k)} \right) \right) \\
&+ \sum_{k=p+1}^q \left(\beta_{2k-1}(n) \sum_{j=g}^{n-1} \Re \left(M_{j+1}^{(k)} - M_j^{(k)} \right) + \beta_{2k}(n) \sum_{j=g}^{n-1} \Im \left(M_{j+1}^{(k)} - M_j^{(k)} \right) \right) + r_n.
\end{aligned}$$

For n fixed, this is an *infinite* martingale difference array plus error term r_n which tends to 0 almost surely as $n \rightarrow \infty$. We now cut off the tail of the series to work with a finite martingale difference array. More precisely, we choose a sequence $(k_n)_{n \geq 0} \uparrow \infty$ appropriately and write

$$\begin{aligned}
\sum_{k=1}^{2q} \alpha_k Z_{n,k} &= \sum_{k=1}^p \left(\beta_{2k-1}(n) \sum_{j=n}^{k_n} \Re \left(M_j^{(k)} - M_{j+1}^{(k)} \right) + \beta_{2k}(n) \sum_{j=n}^{k_n} \Im \left(M_j^{(k)} - M_{j+1}^{(k)} \right) \right) \\
&+ \sum_{k=p+1}^q \left(\beta_{2k-1}(n) \sum_{j=g}^{n-1} \Re \left(M_{j+1}^{(k)} - M_j^{(k)} \right) + \beta_{2k}(n) \sum_{j=g}^{n-1} \Im \left(M_{j+1}^{(k)} - M_j^{(k)} \right) \right) + \varepsilon_n
\end{aligned}$$

such that $\varepsilon_n \rightarrow 0$ in L^2 . The following lemma shows that $(k_n)_{n \geq 0} = (n^2)_{n \geq 0}$ is sufficient.

Lemma 2.1. *Let*

$$\varepsilon_n := \sum_{k=1}^p \left(\beta_{2k-1}(n) \sum_{j=n^2+1}^{\infty} \Re \left(M_j^{(k)} - M_{j+1}^{(k)} \right) + \beta_{2k}(n) \sum_{j=n^2+1}^{\infty} \Im \left(M_j^{(k)} - M_{j+1}^{(k)} \right) \right) + r_n.$$

Then

$$\varepsilon_n \xrightarrow{L^2} 0$$

as $n \rightarrow \infty$.

Proof. It is easy to see that r_n tends to zero in L^2 as $\Re(\lambda_k) < r/2$ for all summands in this term. The remaining part follows immediately from Lemma 1.3. \square

For the proof of Theorem 2.1, it is thus sufficient to show weak convergence of the martingale difference array

$$\begin{aligned} & \sum_{k=1}^p \left(\beta_{2k-1}(n) \sum_{j=n}^{n^2} \Re \left(M_j^{(k)} - M_{j+1}^{(k)} \right) + \beta_{2k}(n) \sum_{j=n}^{n^2} \Im \left(M_j^{(k)} - M_{j+1}^{(k)} \right) \right) \\ & + \sum_{k=p+1}^q \left(\beta_{2k-1}(n) \sum_{j=g}^{n-1} \Re \left(M_{j+1}^{(k)} - M_j^{(k)} \right) + \beta_{2k}(n) \sum_{j=g}^{n-1} \Im \left(M_{j+1}^{(k)} - M_j^{(k)} \right) \right) \end{aligned}$$

which perfectly fits into the setting of Proposition 2.2. We now check the conditions.

Depending on the summation index j , there are two types of increments $I_{n,j}$. We use the shorthand

$$I_{n,j} := \begin{cases} \sum_{k=p+1}^q \left(\beta_{2k-1}(n) \Re \left(M_{j+1}^{(k)} - M_j^{(k)} \right) + \beta_{2k}(n) \Im \left(M_{j+1}^{(k)} - M_j^{(k)} \right) \right), & j < n, \\ \sum_{k=1}^p \left(\beta_{2k-1}(n) \Re \left(M_j^{(k)} - M_{j+1}^{(k)} \right) + \beta_{2k}(n) \Im \left(M_j^{(k)} - M_{j+1}^{(k)} \right) \right), & j \geq n. \end{cases}$$

The absolute value of these increments is deterministically bounded: For $j < n$,

$$\begin{aligned} |I_{n,j}| & \leq \sum_{k=p+1}^q |\beta_{2k-1}(n)| \left| \Re \left(M_{j+1}^{(k)} - M_j^{(k)} \right) \right| + |\beta_{2k}(n)| \left| \Im \left(M_{j+1}^{(k)} - M_j^{(k)} \right) \right| \\ & \leq C \sum_{k=p+1}^q n^{\Re(\lambda_k)/r-1/2} \left(\left| \Re(M_{j+1}^{(k)} - M_j^{(k)}) \right| + \left| \Im(M_{j+1}^{(k)} - M_j^{(k)}) \right| \right) \\ & \leq \sqrt{2}C \sum_{k=p+1}^q n^{\Re(\lambda_k)/r-1/2} \left| M_{j+1}^{(k)} - M_j^{(k)} \right| \\ & \leq D \ n^{-1/2} \sum_{k=p+1}^q \left(\frac{n}{j} \right)^{\Re(\lambda_k)/r} = O \left(n^{\max\{\Re(\lambda_{p+1})/r, 0\}-1/2} \right) \end{aligned}$$

for all $j < n$. Here, $C, D > 0$ are constants and the uniform bound tends to 0 as $n \rightarrow \infty$. Analogously, for $n \leq j \leq n^2$,

$$|I_{n,j}| \leq C \ n^{-1/2} \sum_{k=1}^p \left(\frac{n}{j} \right)^{\Re(\lambda_k)/r} = O \left(n^{-1/2} \right)$$

for all $n \leq j \leq n^2$ as $n \rightarrow \infty$. Again, $C > 0$ is a constant and the uniform bound tends to 0 as $n \rightarrow \infty$.

Fix $\varepsilon > 0$. Then by the above, for N big enough, for all $j = 1, \dots, N^2$ we have $|I_{n,j}| < \varepsilon$, so in particular

$$\sum_{j=g}^{n^2} \mathbb{E}[I_{n,j}^2 1(|I_{n,j}| > \varepsilon) | \mathcal{F}_{n,j-1}] = 0$$

for all $n \geq N$. This implies condition (14).

We now turn to condition (15), which is computationally the hardest.

We rewrite the increments as

$$I_{n,j}^2 = \left(\xi_{n,j}^t (X_{j+1} - X_j) + \eta_{n,j}^t \frac{X_j}{rj + |X_0|} \right)^2$$

where

$$\begin{aligned} \xi_{n,j} := & \sum_{k \in K} \frac{1}{\sqrt{n}} \left[\left(\alpha_{2k-1} \Re \left(\frac{\gamma_n^{(k)}}{\gamma_{j+1}^{(k)}} \right) + \alpha_{2k} \Im \left(\frac{\gamma_n^{(k)}}{\gamma_{j+1}^{(k)}} \right) \right) \Re(\bar{u}_k) \right. \\ & \left. + \left(\alpha_{2k} \Re \left(\frac{\gamma_n^{(k)}}{\gamma_{j+1}^{(k)}} \right) - \alpha_{2k-1} \Im \left(\frac{\gamma_n^{(k)}}{\gamma_{j+1}^{(k)}} \right) \right) \Im(\bar{u}_k) \right] \end{aligned}$$

and

$$\begin{aligned} \eta_{n,j} := & \sum_{k \in K} \frac{1}{\sqrt{n}} \left[\left(\alpha_{2k-1} \Re \left(\frac{\gamma_n^{(k)}}{\gamma_{j+1}^{(k)}} \right) + \alpha_{2k} \Im \left(\frac{\gamma_n^{(k)}}{\gamma_{j+1}^{(k)}} \right) \right) \Re(\bar{\lambda}_k \bar{u}_k) \right. \\ & \left. + \left(\alpha_{2k} \Re \left(\frac{\gamma_n^{(k)}}{\gamma_{j+1}^{(k)}} \right) - \alpha_{2k-1} \Im \left(\frac{\gamma_n^{(k)}}{\gamma_{j+1}^{(k)}} \right) \right) \Im(\bar{\lambda}_k \bar{u}_k) \right] \end{aligned}$$

and the set K is either equal to $\{1, \dots, p\}$ or $\{p+1, \dots, q\}$, depending on j . With this,

$$\begin{aligned} \sum_{j=g}^{n^2} \mathbb{E}[I_{n,j}^2 | \mathcal{F}_{n,j-1}] &= \sum_{j=g}^{n^2} \mathbb{E} \left[\left(\xi_{n,j}^t (X_{j+1} - X_j) - \eta_{n,j}^t \frac{X_j}{rj + |X_0|} \right)^2 | \mathcal{F}_{j-1} \right] \\ &= \sum_{j=g}^{n^2} \sum_{m=1}^q \frac{X_j^{(m)}}{rj + |X_0|} \left(\xi_{n,j}^t \Delta_m - \eta_{n,j}^t \frac{X_j}{rj + |X_0|} \right)^2. \end{aligned} \quad (16)$$

We show that this converges almost surely by looking at the different terms separately: Recall that each of the $\xi_{n,j}$ and $\eta_{n,j}$ is a sum over different eigenspace components.

In the first part of the sum where $j \leq n-1$, there are only λ_k with $\Re(\lambda_k)/r < 1/2$, so the product of component k with component ℓ in the square $\left(\xi_{n,j}^t \Delta_m - \eta_{n,j}^t \frac{X_j}{rj + |X_0|} \right)^2$ is asymptotically equivalent to

$$\begin{aligned} & \frac{1}{n} \sum_{j=g}^{n-1} \sum_{m=1}^q \frac{X_j^{(m)}}{rj + |X_0|} \left(\frac{n}{j} \right)^{\frac{\Re(\lambda_k + \lambda_\ell)}{r}} \\ & \cdot \left(\left(\alpha_{2k-1} \Re \left(\bar{\lambda}_k \left(\bar{u}_k^{(m)} - \frac{\bar{u}_k^t X_j}{rj + |X_0|} \right) \right) + \alpha_{2k} \Im \left(\bar{\lambda}_k \left(\bar{u}_k^{(m)} - \frac{\bar{u}_k^t X_j}{rj + |X_0|} \right) \right) \right) \cos \left(\frac{\Im(\bar{\lambda}_k)}{r} \log \left(\frac{n}{j} \right) \right) \right. \\ & \left. + \left(\alpha_{2k} \Re \left(\bar{\lambda}_k \left(\bar{u}_k^{(m)} - \frac{\bar{u}_k^t X_j}{rj + |X_0|} \right) \right) - \alpha_{2k-1} \Im \left(\bar{\lambda}_k \left(\bar{u}_k^{(m)} - \frac{\bar{u}_k^t X_j}{rj + |X_0|} \right) \right) \right) \sin \left(\frac{\Im(\bar{\lambda}_k)}{r} \log \left(\frac{n}{j} \right) \right) \right) \\ & \cdot \left(\left(\alpha_{2\ell-1} \Re \left(\bar{\lambda}_\ell \left(\bar{u}_\ell^{(m)} - \frac{\bar{u}_\ell^t X_j}{rj + |X_0|} \right) \right) + \alpha_{2\ell} \Im \left(\bar{\lambda}_\ell \left(\bar{u}_\ell^{(m)} - \frac{\bar{u}_\ell^t X_j}{rj + |X_0|} \right) \right) \right) \cos \left(\frac{\Im(\bar{\lambda}_\ell)}{r} \log \left(\frac{n}{j} \right) \right) \right. \\ & \left. + \left(\alpha_{2\ell} \Re \left(\bar{\lambda}_\ell \left(\bar{u}_\ell^{(m)} - \frac{\bar{u}_\ell^t X_j}{rj + |X_0|} \right) \right) - \alpha_{2\ell-1} \Im \left(\bar{\lambda}_\ell \left(\bar{u}_\ell^{(m)} - \frac{\bar{u}_\ell^t X_j}{rj + |X_0|} \right) \right) \right) \sin \left(\frac{\Im(\bar{\lambda}_\ell)}{r} \log \left(\frac{n}{j} \right) \right) \right). \end{aligned}$$

The vector $\frac{X_j}{rj+|X_0|}$ converges almost surely to V . It follows that for $\lambda_k \neq r$, $\bar{u}_k^t \frac{X_j}{rj+|X_0|} \rightarrow 0$ almost surely. So the sum above is almost surely asymptotically equivalent to

$$\begin{aligned}
& \frac{1}{2n} \sum_{j=g}^{n-1} \sum_{m=1}^q V^{(m)} \left(\frac{n}{j} \right)^{\frac{\Re(\lambda_k + \lambda_\ell)}{r}} \\
& \cdot \left(((\alpha_{2k-1} \alpha_{2\ell-1} - \alpha_{2k} \alpha_{2\ell}) \Re(\bar{\lambda}_k \bar{\lambda}_\ell \bar{u}_k^{(m)} \bar{u}_\ell^{(m)}) + (\alpha_{2k-1} \alpha_{2\ell} + \alpha_{2k} \alpha_{2\ell-1}) \Im(\bar{\lambda}_k \bar{\lambda}_\ell \bar{u}_k^{(m)} \bar{u}_\ell^{(m)})) \right) \\
& \cdot \cos \left(\frac{\Im(\bar{\lambda}_k + \bar{\lambda}_\ell)}{r} \log \left(\frac{n}{j} \right) \right) + ((\alpha_{2k-1} \alpha_{2\ell-1} + \alpha_{2k} \alpha_{2\ell}) \Re(\lambda_k \bar{\lambda}_\ell u_k^{(m)} \bar{u}_\ell^{(m)}) \\
& + (\alpha_{2k-1} \alpha_{2\ell} - \alpha_{2k} \alpha_{2\ell-1}) \Im(\lambda_k \bar{\lambda}_\ell u_k^{(m)} \bar{u}_\ell^{(m)})) \cos \left(\frac{\Im(\bar{\lambda}_k - \bar{\lambda}_\ell)}{r} \log \left(\frac{n}{j} \right) \right) \\
& + ((\alpha_{2k-1} \alpha_{2\ell} + \alpha_{2k} \alpha_{2\ell-1}) \Re(\bar{\lambda}_k \bar{\lambda}_\ell \bar{u}_k^{(m)} \bar{u}_\ell^{(m)}) + (-\alpha_{2k-1} \alpha_{2\ell-1} + \alpha_{2k} \alpha_{2\ell}) \Im(\bar{\lambda}_k \bar{\lambda}_\ell \bar{u}_k^{(m)} \bar{u}_\ell^{(m)})) \\
& \cdot \sin \left(\frac{\Im(\bar{\lambda}_k + \bar{\lambda}_\ell)}{r} \log \left(\frac{n}{j} \right) \right) + ((-\alpha_{2k-1} \alpha_{2\ell} + \alpha_{2k} \alpha_{2\ell-1}) \Re(\lambda_k \bar{\lambda}_\ell u_k^{(m)} \bar{u}_\ell^{(m)}) \\
& + (\alpha_{2k-1} \alpha_{2\ell-1} + \alpha_{2k} \alpha_{2\ell}) \Im(\lambda_k \bar{\lambda}_\ell u_k^{(m)} \bar{u}_\ell^{(m)})) \sin \left(\frac{\Im(\bar{\lambda}_k - \bar{\lambda}_\ell)}{r} \log \left(\frac{n}{j} \right) \right) \\
& \rightarrow \alpha_{2k-1} \alpha_{2\ell-1} (\Sigma_V)_{2k-1, 2\ell-1} + \alpha_{2k} \alpha_{2\ell} (\Sigma_V)_{2k, 2\ell} + \alpha_{2k-1} \alpha_{2\ell} (\Sigma_V)_{2k-1, 2\ell} + \alpha_{2k} \alpha_{2\ell-1} (\Sigma_V)_{2k, 2\ell-1}
\end{aligned}$$

as $n \rightarrow \infty$.

Essentially the same calculation shows convergence for summands k, ℓ with $1/2 < \Re(\lambda_k)/r, \Re(\lambda_\ell)/r < 1$. Furthermore, there are no summands with $\Re(\lambda_k)/r \leq 1/2$ and $\Re(\lambda_\ell)/r > 1/2$. Let now $1/2 < \Re(\lambda_k)/r < 1$ and $\lambda_\ell = r$. Then the component k with component ℓ product converges to

$$\begin{aligned}
& \alpha_{2k-1} \alpha_{2\ell-1} \cdot r^2 \sum_{m=1}^q V^{(m)} u_\ell^{(m)} \Re(u_k^{(m)}) + \alpha_{2k} \alpha_{2\ell-1} \cdot r^2 \sum_{m=1}^q V^{(m)} u_\ell^{(m)} \Im(u_k^{(m)}) \\
& = 0 \\
& = \alpha_{2k-1} \alpha_{2\ell-1} (\Sigma_V)_{2k-1, 2\ell-1} \alpha_{2k-1} \alpha_{2\ell} (\Sigma_V)_{2k-1, 2\ell}.
\end{aligned}$$

Finally, if $\lambda_k = \lambda_\ell = r$, the product tends to

$$\begin{aligned}
& \alpha_{2k-1} \alpha_{2\ell-1} r^2 \sum_{m=1}^q V^{(m)} \left(u_k^{(m)} - \left(\frac{\Xi_k}{r} + \frac{\pi_k(X_0)}{|X_0|} \right) \right) \left(u_\ell^{(m)} - \left(\frac{\Xi_\ell}{r} + \frac{\pi_\ell(X_0)}{|X_0|} \right) \right) \\
& = \begin{cases} -\alpha_{2k-1} \alpha_{2\ell-1} r^2 \left(\frac{\Xi_k}{r} + \frac{\pi_k(X_0)}{|X_0|} \right) \left(\frac{\Xi_\ell}{r} + \frac{\pi_\ell(X_0)}{|X_0|} \right), & k \neq \ell \\ \alpha_{2k-1}^2 r^2 \left(\frac{\Xi_k}{r} + \frac{\pi_k(X_0)}{|X_0|} \right) \left(1 - \left(\frac{\Xi_k}{r} + \frac{\pi_k(X_0)}{|X_0|} \right) \right), & k = \ell \end{cases} \\
& = \alpha_{2k-1} \alpha_{2\ell-1} (\Sigma_V)_{2k-1, 2\ell-1}.
\end{aligned}$$

In total, this implies that

$$\sum_{j=g}^{n^2} \mathbb{E}[I_{n,j}^2 | \mathcal{F}_{n,j-1}] \xrightarrow{a.s.} \sum_{i,j=1}^{2q} \alpha_i \alpha_j (\Sigma_V)_{i,j} = (\alpha_1, \dots, \alpha_{2q}) \Sigma_V (\alpha_1, \dots, \alpha_{2q})^t.$$

By Proposition 2.2, $\alpha_1 Z_n^{(1)} + \dots + \alpha_{2q} Z_n^{(2q)}$ converges weakly to a random variable with characteristic function $\mathbb{E}[\exp(-1/2((\alpha_1, \dots, \alpha_{2q}) \Sigma_V (\alpha_1, \dots, \alpha_{2q})^t)^2 t^2)]$. Now the Cramér-Wold device implies weak convergence of $(Z_n)_n$ to a mixed multivariate normal distribution with covariance matrix Σ_V given V : If the almost sure limit V is deterministic, then the asymptotic distribution of Z_n is the normal distribution, as can be seen from the characteristic function. If the almost sure limit V is not deterministic, then, conditionally on V , Z_n converges to some normal limit law, so the asymptotic distribution of Z_n is a mixed normal distribution.

We finally consider the case where there is some k such that $\Re(\lambda_k)/r = 1/2$. By very similar calculations, there is also weak convergence of $\alpha_1 Z_n^{(1)} + \dots + \alpha_{2q} Z_n^{(2q)}$ to a random variable with characteristic function $\mathbb{E}[\exp(-1/2((\alpha_1, \dots, \alpha_{2q}) \Sigma_V (\alpha_1, \dots, \alpha_{2q})^t)^2 t^2)]$, where Σ_V is defined in equation (12). Due to the scaling, the matrix Σ_V has a lot more zero entries in this case. Again, the Cramér-Wold device implies Theorem 2.1 in this case. \square

Proof of Theorem 1.3. It remains to show that under our assumptions, in both cases of Theorem (1.3), the covariance matrix A_V given V is of the stated form and almost surely has positive entries in the specified positions. Because all complex conjugates of eigenvectors in our basis are also eigenvectors in the basis, we have the almost sure asymptotic equivalence

$$\frac{1}{\sqrt{n\ell_n}} \left(Y_n - \sum_{k=1}^p n^{\frac{\lambda_k}{r}} \Xi_k v_k \right) \sim \sum_{k=1}^q \left(Z_n^{(2k-1)} \Re(v_k) - Z_n^{(2k)} \Im(v_k) \right),$$

where $\ell_n = 1$ in the first case and $\ell_n = \log(n)$ in the second case. This shows that $A_V = M \Sigma_V M^t$ in both cases.

The components on the right hand side are linear combinations of the components of Z_n as considered in the last proof. For a fixed colour j , the conditional variance of colour j is the conditional variance of

$$\sum_{k=1}^q \left(Z_n^{(2k-1)} \Re(v_k^{(j)}) - Z_n^{(2k)} \Im(v_k^{(j)}) \right). \quad (17)$$

First case: Non-dominant colours. If j is a non-dominant colour, $v_k^{(j)} = 0$ for all dominant colours k by our choice of right eigenvectors. So (17) reduces to a sum over type 3 colours. But all variances and covariances of type 3 projections are zero in the limit and $(A_V)_{j,j} = 0$.

Second case: Dominant colours. Suppose that j is a dominant colour in class \mathcal{C}_m . Again, by our choice of right eigenvectors, the sum reduces to a sum over colours in \mathcal{C}_m , as $v_k^{(j)} \neq 0$ only if k is a colour in class \mathcal{C}_m . If $|\mathcal{C}_m| = 1$, then

$$(A_V)_{j,j} = r^2 \left(\frac{\Xi_m}{r} + \frac{\pi_k(X_0)}{|X_0|} \right) \left(1 - \left(\frac{\Xi_m}{r} + \frac{\pi_k(X_0)}{|X_0|} \right) \right) > 0$$

almost surely by Remark 4, because there is at least one other dominant class by our assumptions. Let $|\mathcal{C}_m| > 1$. We now look at the different components of the variance of colour j coming from projections associated with class \mathcal{C}_m in more detail. Remember that with coefficients chosen appropriately, the sum (16) converges to the conditional variance of (17) almost surely as $n \rightarrow \infty$. Each of its $n^2 + 1 - g$ summands is non-negative and thus any sum over less terms yields a lower bound for the whole sum. Only considering part of the sum has the advantage that the variances and covariances of the fluctuations in the sum grow at a different speed in the beginning (respectively end) of the urn. For example, for n large and $\varepsilon \in (0, 1)$, we can either sum from g to εn or from $\varepsilon^{-1}n$ to n^2 to get a lower bound. In the first case, there are only summands with $\Re(\lambda_k), \Re(\lambda_\ell) \leq r/2$. A calculation as in the proof of Theorem 2.1 shows that the contribution coming from the fluctuations in projections π_k, π_ℓ to the sum (16) cut off at εn with coefficients chosen appropriately is at most of order $\varepsilon^{1-\Re(\lambda_k+\lambda_\ell)/r}$. In the second case, there are only summands with $\Re(\lambda_k), \Re(\lambda_\ell) > r/2$. The contribution coming from the fluctuations in projections π_k, π_ℓ to the sum (16) without the first $\varepsilon^{-1}n$ summands with coefficients chosen appropriately is at most of order $\varepsilon^{\Re(\lambda_k+\lambda_\ell)/r-1}$. In particular, the variance contribution from projections with real part close to $r/2$ (and nonzero coefficients) is the greatest.

We now choose k such that among all possible $\lambda_k \neq 0$ associated to \mathcal{C}_m , the distance $|\Re(\lambda_k)/r - 1/2|$ is minimal and $|v_k^{(j)}| > 0$. Note that this is possible as the Perron-Frobenius eigenvalue associated with \mathcal{C}_m satisfies these conditions, for example.

In case 1 of Theorem 1.3, there are only the cases $\Re(\lambda_k)/r < 1/2$ and $\Re(\lambda_k)/r > 1/2$. Assume

$\Re(\lambda_k)/r < 1/2$. If $\lambda_k \in \mathbb{R}$ is a simple eigenvalue, there is only one dominant term of order

$$\frac{|\lambda_k|^2 |\mathbf{v}_k^{(j)}|^2}{\Re(1 - 2\lambda_k/r)} \sum \mathbf{V}^{(m)} |\mathbf{u}_k^{(m)}|^2,$$

which is positive. If $\lambda_k \in \mathbb{C} \setminus \mathbb{R}$ is a simple eigenvalue, we choose λ_k with $\Re(\lambda_k) > 0$ and ε small enough such that $2\Im(\lambda_k)/r \log(\varepsilon)$ is a negative multiple of 2π . We cut off at $\varepsilon \mathbf{n}$. The dominant term is of order $\varepsilon^{1-2\Re(\lambda_k)/r}$. It has coefficient

$$2 \frac{|\lambda_k|^2 |\mathbf{v}_k^{(j)}|^2}{\Re(1 - 2\lambda_k/r)} \sum \mathbf{V}^{(m)} |\mathbf{u}_k^{(m)}|^2 + \frac{2}{|1 - 2\lambda_k/r|^2} \sum \mathbf{V}^{(m)} \operatorname{Re} \left((1 - 2\bar{\lambda}_k/r) \lambda_k^2 (\mathbf{u}_k^{(m)})^2 (\bar{\mathbf{v}}_k^{(j)})^2 \right),$$

which is positive (remember $|\mathbf{v}_k^{(j)}| > 0$). Now if λ_k is a multiple eigenvalue with these properties, the covariances of the associated components of \mathbf{Z} are zero due to our choice of multiple eigenvalues in Section 1.2, and hence each projection associated to this eigenvalue yields a lower bound on the variance.

If $\Re(\lambda_k)/r > 1/2$, we choose ε small enough, start the sum at $\varepsilon^{-1} \mathbf{n}$ and proceed analogously. The dominant term now is of order $\varepsilon^{2\Re(\lambda_k)/r-1}$ and has non-zero coefficient.

If there are a big eigenvalue λ_k and a small eigenvalue λ_ℓ such that $|\Re(\lambda_k)/r - 1/2| = |\Re(\lambda_\ell)/r - 1/2|$, we can choose either.

In case 2 of Theorem 1.3, there are k in class \mathcal{C}_m such that $\Re(\lambda_k)/r = 1/2$. This case is even simpler as everything that is not killed by the scaling is independent. The asymptotic variance of colour j is then the weighted sum over all variances of real and imaginary parts of the $r/2$ -projections in the class, which are not identically zero, because all different summands are asymptotically uncorrelated. □

3 Applications

Below, we consider four examples which illustrate the statement of Theorem 1.3 for both small and large urns. Interpreting the first example as a special case of the second, a phase transition takes place within each of the models. Note however, that the result for small urns is covered in [20] and [8].

3.1 Pólya urn

We consider the original Pólya urn with $q = 2$ colours and matrix $\mathbf{R} = \operatorname{Id}_2$. This matrix has $r = 1$ as a multiple eigenvalue, so we expect weak convergence of the error term to a mixed normal distribution. We choose

$$\mathbf{u}_1 := \mathbf{v}_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 := \mathbf{v}_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

as dual bases of left and right eigenvectors.

Note that, for example, $\mathbf{X}_0 = (1, 0)^t$ is no admissible initial configuration in the setting of the paper, as in this case, $\mathbf{X}_n = (n+1)(1, 0)^t$ for all n almost surely, so the probability to have balls of colour 2 at time n is zero for all n and \mathbf{Y}_n is identically $(0, 0)^t$.

But $\mathbf{X}_0 = (1, 1)^t$ is an admissible initial configuration, for example. For this initial configuration, the limit \mathbf{V} is given by $\mathbf{V} = (\mathbf{U}, 1 - \mathbf{U})^t$, where \mathbf{U} is uniformly distributed on $(0, 1)$. By explicitly calculating the variance in Theorem 1.3, we recover a result of Hall and Heyde [7] p. 80,

$$\frac{1}{\sqrt{n}} \left(\mathbf{Y}_n - n \begin{pmatrix} \mathbf{U} - \frac{1}{2} \\ \frac{1}{2} - \mathbf{U} \end{pmatrix} \right) \xrightarrow{\mathcal{L}} \sqrt{\mathbf{U}(1 - \mathbf{U})} \cdot \mathcal{N} \left(0, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \right),$$

where \mathbf{U} is independent of the normal random vector \mathcal{N} . Note that because the almost sure limit of $\mathbf{X}_n/(n+2)$ is random, a mixed normal distribution arises.

3.2 Friedman's urn

As an extension of the previous example, consider Friedman's urn with generating matrix

$$R = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix},$$

where $\alpha, \beta \in \mathbb{Z}$, $\alpha \geq -1$, $\beta \geq 0$, and $\alpha + \beta = r > 0$. This symmetric matrix has real eigenvalues $\lambda_1 := \alpha + \beta$ and $\lambda_2 := \alpha - \beta$. For $\beta = 0$, we get the original Pólya urn which can be treated as the previous example. For $\beta > 0$, $\lambda_1 = r$ is a simple eigenvalue and thus no mixing over normal distributions arises in our limit theorems. We choose eigenvectors

$$u_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad u_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad v_1 := \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_2 := \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

For $\beta > 0$, the behaviour of the urn ranges over three cases: If $\alpha < 3\beta$, then $\lambda_2 < \lambda_1/2$, and

$$\begin{aligned} A_V &= M \Sigma_V M^t = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} & 0 \end{pmatrix} \frac{(\alpha + \beta)(\alpha - \beta)^2}{3\beta - \alpha} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 \end{pmatrix} \\ &= \frac{(\alpha + \beta)(\alpha - \beta)^2}{4(3\beta - \alpha)} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \end{aligned}$$

so

$$\frac{Y_n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{(\alpha + \beta)(\alpha - \beta)^2}{4(3\beta - \alpha)} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\right)$$

as $n \rightarrow \infty$.

If $\alpha = 3\beta$, then $\lambda_2 = \lambda_1/2$, a very similar calculation leads to

$$\frac{Y_n}{\sqrt{n \log(n)}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{(\alpha - \beta)^2}{4} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\right),$$

as $n \rightarrow \infty$. Here, we have recovered the results of Example 3.27 in [8].

If $\alpha > 3\beta$, then $\lambda_2 > \lambda_1/2$, and so

$$\frac{Y_n - n^{\frac{\alpha - \beta}{\alpha + \beta}} \Xi_2 v_2}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{(\alpha + \beta)(\alpha - \beta)^2}{4(\alpha - 3\beta)} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\right)$$

as $n \rightarrow \infty$. Ξ_2 is the almost sure limit of the martingale $\Gamma\left(\frac{|X_0|}{\alpha + \beta}\right) \Gamma\left(\frac{|X_0| + \alpha - \beta}{\alpha + \beta}\right)^{-1} (\gamma_n^{(2)})^{-1} (u_2^t Y_n)$.

3.3 m-ary search tree

The evolution of the vector of the number of nodes containing $0, 1, \dots, m-2$ keys in an m -ary search tree under the uniform permutation model can be encoded by the following urn model: We have generating matrix

$$R_m = \begin{pmatrix} -1 & 0 & & & m \\ 2 & -2 & & & \\ & 3 & -3 & & \\ & & & \ddots & \\ & & & m-1 & -(m-1) \end{pmatrix}$$

and $X_0 = (1, 0, \dots, 0)^t$. It is well known, cf. [3], that

$$\frac{X_n}{n+1} \xrightarrow{\text{a.s.}} \frac{1}{H_m - 1} \left(\frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{m} \right)^t$$

as $n \rightarrow \infty$, where H_m denotes the m -th Harmonic number. This limit is deterministic and all its components are strictly positive, so we expect convergence in distribution to a non-mixed Gaussian distribution. The eigenvalues of R_m are given by the solutions of the equation

$$m! = \prod_{k=1}^{m-1} (z + k).$$

If $m \leq 26$, there are no eigenvalues with real part greater than $1/2$. In this case, Theorem 1.3 confirms the well-known result that $(X_n - \mathbb{E}[X_n])/\sqrt{n} \rightarrow \mathcal{N}$ in distribution: Mahmoud and Pittel [15] showed that when $m \leq 15$, the limiting distribution is normal. The result was later extended to include $m \leq 26$ by Lew and Mahmoud [13].

For $m > 26$, there is at least one eigenvalue with real part greater than $1/2$ and it is known that for all such m , there is no eigenvalue whose real part is equal to $1/2$. Chern and Hwang [5] proved that when $m \geq 27$, the space requirement centered by its mean and scaled by its standard deviation does not have a limiting distribution. Here, Theorem 1.3 can be applied to give the result

$$\frac{1}{\sqrt{n}} \left(X_n - \mathbb{E}[X_n] - \sum_{k=1}^p n^{\lambda_k} \Xi_k v_k \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A_V).$$

3.4 B-urns

A B-tree with integer-valued parameter $m \geq 2$ is a search tree, where keys are stored in internal nodes of the tree and its leaves (or gaps) represent future insertion possibilities. See [4] for a nice exposition of B-trees. When constructed from the so-called optimistic algorithm (see [4]), there are $m - 1$ different types of gaps. As explored in [4], the evolution of the gap process $(X_n)_{n \geq 0}$ - i.e., the joint evolution of the gaps of different types - under the random permutation model on the keys can be modelled as a Pólya urn with generating matrix

$$R_m = \begin{pmatrix} -m & 0 & & & 2m \\ m+1 & -(m+1) & & & \\ & m+2 & -(m+3) & & \\ & & & \ddots & \\ & & & 2m-1 & -(2m-1) \end{pmatrix}.$$

The eigenvalues of R_m are given by the solutions of the equation

$$\frac{2m!}{m!} = \prod_{k=m}^{2m-1} (z + k).$$

Its complex roots are all simple, and 1 is the Perron Frobenius eigenvalue. Furthermore, two distinct eigenvalues that have the same real part are conjugated. The left and right eigenvectors of R_m are also explicitly calculated in [4].

If $m \leq 59$, there are no eigenvalues with real part greater than $1/2$. In this case, Theorem 1.3 confirms that

$$\frac{X_n - \mathbb{E}[X_n]}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, A_V).$$

For $m \geq 60$, Theorem 1.3 states that

$$\frac{1}{\sqrt{n}} \left(X_n - \mathbb{E}[X_n] - \sum_{k=1}^p n^{\lambda_k} \Xi_k v_k \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A_V).$$

References

- [1] G. K. Basak and A. Dasgupta. Central limit theorems for a class of irreducible multicolor urn models. *Proc. Indian Acad. Sci. (Math. Sci.)*, Vol. 117, No. 4 (2005), 517 – 543.
- [2] P. Berti, I. Crimaldi, L. Pratelli and P. A. Rigo. Central Limit Theorem and its Applications to Multicolour Randomly Reinforced Urns. *Journal of Applied Probability* Vol. 48 No. 2 (2011), 527 – 546.
- [3] B. Chauvin, and N. Pouyanne. m -ary search trees when $m \geq 27$: a strong asymptotics for the space requirements. *Random Structures and Algorithms*, 24 (2004), 133 – 154.
- [4] B. Chauvin, D. Gardy, N. Pouyanne and D.-H. Ton-That. B-urns. *ALEA, Lat. Am. J. Probab. Math. Stat.* 13 (2016), 605 – 634.
- [5] Chern, H.-H., and Hwang, H.-K., Phase changes in random m -ary search trees and generalized quicksort. *Random Structures Algorithms*, 19 (2001) : 316 – 358.
- [6] Gouet, R., Strong Convergence of Proportions in a Multicolor Pólya Urn. *J. Appl. Prob.*, 34 (1997), 426 – 435.
- [7] Hall, P., and Heyde, C., Martingale Limit Theory and Applications. *Academic Press, New York* (1980).
- [8] Janson, S., Functional limit theorems for multitype branching processes and generalised Pólya urns. *Stoch. Process. Appl.* 110 (2004), 177 – 245.
- [9] Janson, S., Mean and Variance of Balanced Pólya Urns. Preprint available via *arXiv: 1602.06203v1* (2016).
- [10] Knape, M., and Neininger, R., Pólya Urns via the Contraction Method. *Combinatorics, Probability & Computing*, 23 (2014), 1148 – 1186.
- [11] Kuba, M., and Sulzbach, H., On martingale tail sums in affine two-color urn models with multiple drawings. Preprint available via *arXiv: 1511.01618* (2015).
- [12] Laruelle, S., and Pagès, G., Randomized urn models revisited using stochastic approximation. *Ann. Appl. Prob.* Volume 23 (2013), 1409-1436.
- [13] Lew, W. and Mahmoud, H. M., *The joint distribution of elastic buckets in multiway search trees*. *SIAM J. Comput.*, 23 (1994):1050 – 1074.
- [14] Mahmoud, H. M., Pólya Urn Models. *CLC Press, Boca Raton, FL* (2009).
- [15] Mahmoud, H. M., and Pittel, B., *Analysis of the space of search trees under the random insertion algorithm*. *J. Algorithms*, 10 (1989): 52 – 75.
- [16] Mueller, N., and Neininger, R., The CLT Analogue for Cyclic Urns. *Analytic Algorithmics and Combinatorics (ANALCO)* (2016), 121 – 127.
- [17] Neininger, R., Refined Quicksort asymptotics. *Random Structures and Algorithms* 46 (2015), 346 – 361.
- [18] Pouyanne, N., An algebraic approach to Pólya processes. *Ann. Inst. H. Poincaré Probab. Statist.* Volume 44, (2008), 293 – 323.
- [19] Seneta, E., Non-negative Matrices and Markov Chains. *Springer, Berlin* (1981).
- [20] Smythe, R. T., Central limit theorems for urn models. *Stochastic Processes and their Applications*, Volume 65 (1996), 115 – 137.
- [21] Sulzbach, H., The law of the iterated logarithm for the path length in random binary trees. Preprint available via *arXiv: 1412.3508* (2015).